PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : G06F 17/30 — A1

(11) International Publication Number: **WO 00/10107**

(43) International Publication Date: 24 February 2000 (24.02.00)

(21) International Application Number: PCT/US99/18645

(22) International Filing Date: 16 August 1999 (16.08.99)

(30) Priority Data:
60/097,029 17 August 1998 (17.08.98) US

(71) Applicant: IATLAS CORPORATION [US/US]; 312 Laurel Avenue, Laurel, MD 20707 (US).

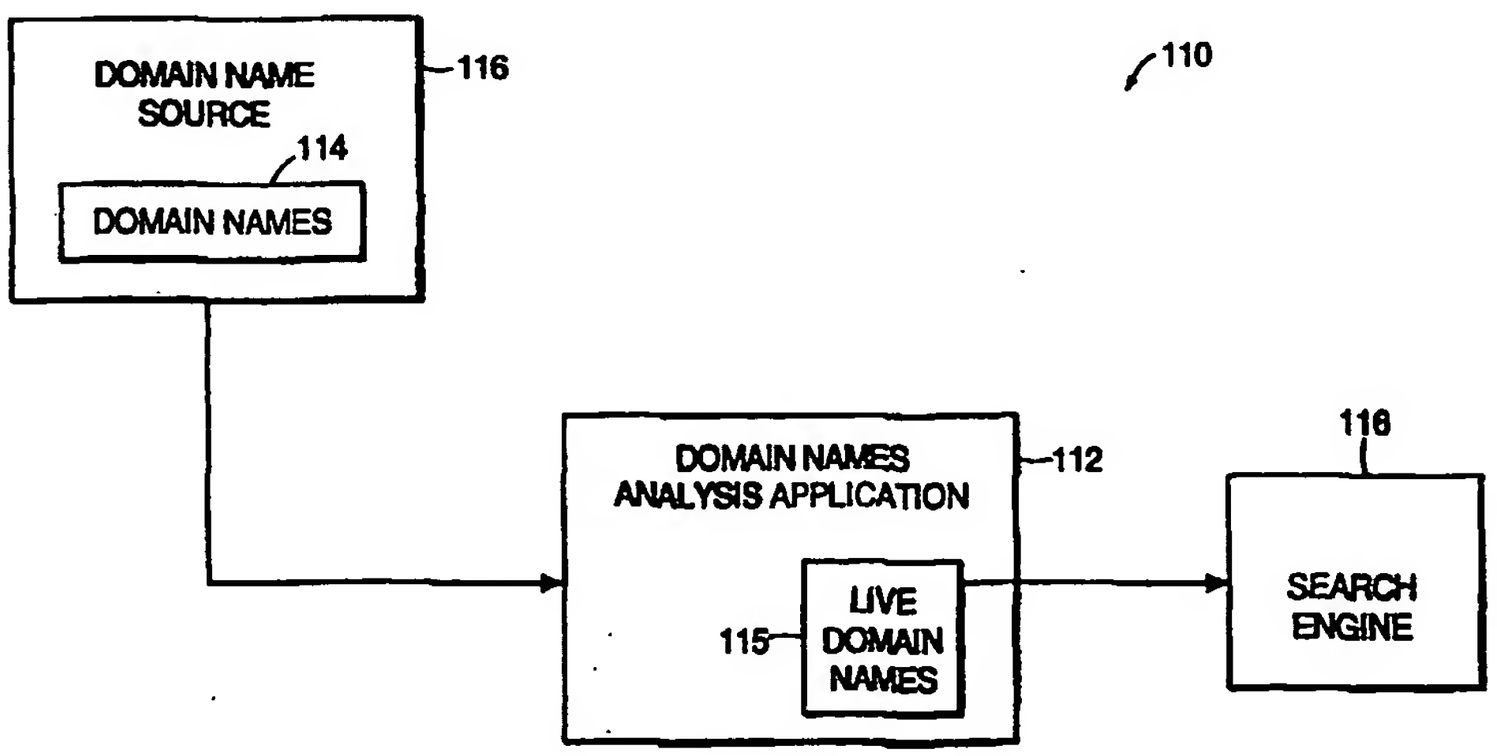
(72) Inventors: BLACK, Jeffrey, Dean; 8706 Bovelder Drive, Laurel, MD 20708 (US). TITUS, Jason, Harvey; 922 Cameron Street, Alexandria, VA 22314 (US). WOODHEAD, Ira, Joseph; 1129 Prince Street, Alexandria, VA 22314 (US).

(74) Agents: REYES, Jason, A. et al.; Hale and Dorr LLP, 60 State Street, Boston, MA 02109 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: ANALYZING INTERNET-BASED INFORMATION



(57) Abstract

A domain name is acquired from a domain name registry. It is determined whether the domain name represents any functioning Web site. If the domain name is associated with at least one functioning Web site, it is recorded that the domain name represents at least one functioning Web site. A set of criteria is acquired. A set of entities is identified that meet the criteria. It is determined how many entities in the set of entities are registered as having control over at least one Web site.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

ANALYZING INTERNET-BASED INFORMATION

Cross-Reference to Related Applications

This application claims the benefit of United States Provisional Application Serial No. 60/097029 entitled "Collecting, Combining, Analyzing, and Using Internet and Business Information" filed on August 17, 1998, which is incorporated herein.

Background of the Invention

This application relates to analyzing Internet-based information.

World-Wide Web ("Web") protocols call for text-based addressing information, which is highly suitable for human users, to be converted to number-based addressing information, which is highly suitable for computers. Much of the information available on the Web is organized into Web pages that can be retrieved and displayed by Web browser software ("browser") under the direction of a user. Each of the Web pages is identifiable by a respective Uniform Resource Locator text string ("URL"), such as "http://www.isp321.com/frontpage.html", that the browser can use to select the page. Each URL includes a domain name, such as "isp321.com", that identifies the Web site where the corresponding Web page is stored for retrieval by browser software. Each domain name is registered by an entity that controls the corresponding Web site and Web pages. A domain name registry organization maintains the domain name registration information,

which may include name, address, and other information that allows the organization to bill the entity for payment for the maintenance. (It is to be understood that the term "registry", as used herein, also refers to a domain name registrar or any other entity that may provide assistance in registering a domain name.) When the domain name is registered, the entity identifies a domain name server computer system that stores a numeric address (known as an IP address) that corresponds to the domain name, and the domain name registry stores the identity of the domain name server computer system together with the domain name in a file known as a zone file. The domain name registry also reports the domain name together with the identity of the domain name server computer system to a root zone server computer system, which is a high level computer system that is responsible for helping other computer systems properly derive IP addresses from domain names (e.g., as described below). The root zone server computer system receives such reports from effectively all domain name registries as domain names are registered, and therefore the root zone server computer system has a comprehensive list of domain names that are registered on the Web.

When the Web browser is directed to retrieve information from a Web site identified by a URL, the browser must determine the IP address of the Web site to which the URL refers. If the browser submits the domain name part of the URL to the root zone server computer system, the root zone server computer system

determines, based on information previously supplied by the domain name registry, the identity of the domain name server computer system for the domain name and reports the identity to the browser. The browser then refers to the domain name server computer system to find out the IP address, and uses the IP address to attempt to contact the Web site. If the Web site is functioning, the browser receives information including a Web page formatting header ("HTML header") from the Web site. If the Web site is not functioning, the browser receives no response from the Web site, and times out indicating an error.

An Internet service provider ("ISP") is an example of an entity that may have a registered domain name for a Web site. Typically, an ISP has customers such as individuals or businesses for whom the ISP stores Web pages on the Web site for retrieval by Web browser software. For example, the ISP may have a customer Maple Street Plumbing for which the ISP stores a home Web page having a URL that includes a prefix "http://www.isp321.com/~maplestplumb". A home Web page is typically the only or the primary entry point into a Web site or a set of Web pages that are under the control of an entity.

Another example of an entity that may have a registered domain name is a Web portal site such as "Yahoo.com" that maintains, in pages organized by categories, links to Web sites and home pages that are under the control of other entities. Typically, a Web portal site allows another entity to create a link from

the Web portal site to the other entity's Web site or home page by submitting information to the Web portal site.

A Web search engine site ("search engine") maintains and updates a search engine database, i.e., a Web page record database, that includes a Web page record for every Web page that has been turned up by Web sweeping software that sweeps the World Wide Web for any and all Web pages. A typical Web page record includes a URL for the respective Web page, an excerpt or other subset of the information provided by the Web page, and a date indicating the most recent update of the Web page record. When a user directs a Web search engine to execute a search, the Web page record database is searched and then search engine results are displayed to the user in the form of a list of Web page records.

The Web sweeping software discovers information on the Web, including domain names previously unknown to the search engine, by following links among Web pages.

Some information about an entity may not be available on a Web site that is under the control of the entity. For example, public financial information about a company may be stored in a database that is not linked to the company's Web site or is not directly accessible by Web browser software, such as a database under the control of a financial services firm.

In general, statistical information regarding Web activity for companies or other entities in a particular sector of human activity, such as an industrial sector, are expressed in broad terms such as the total number of uniquely qualified domain name Web sites ("unique Web sites") that are sponsored by all of the companies in a particular industrial sector. Such statistics may prove misleading for at least some purposes. For example, if ten companies belong to a sector that is known to have ten unique Web sites, the resulting average, i.e., one unique Web site per company, can make the sector appear to be well represented by unique Web sites, even if in actuality all ten unique Web sites belong to only one of the companies and none of the other nine companies has a unique Web site.

Summary of the Invention

Methods and systems are provided for analyzing Internet-based information. An Internet analysis system is provided that gathers domain names and determines whether the domain names are associated with functioning Web sites. A variation of the Internet analysis system that includes an entity information database and a mapping database is able to generate reports regarding Web activity in sectors of human activity such as industrial sectors.

Different aspects of the invention allow one or more of the following. A comprehensive list of tested domain names can be produced. Domain names for

Web sites that are difficult or impossible for a search engine to discover can be made available to the search engine to allow the search engine to produce search results that account for the contents of the previously undiscovered Web sites. Before being provided to the search engine, the domain names may be prioritized or sorted according to one or more attributes (such as industry sector or company size) of the respective entities that are registered as having control over the domain names. Highly useful statistics can be produced concerning the number of entities in an industrial sector that are registered as having control over Web sites. Such statistics can be used for highly effective marketing or sales approaches in which Web oriented products are targeted at potential customers in industrial sectors that are shown by the statistics to have substantial Web activity.

Other features and advantages will become apparent from the following description, including the drawings, and from the claims.

Brief Description of the Drawings

Figs. 1, 6, and 7 are block diagrams of computer-based systems.

Figs. 2, 3, 4, 5, and 10 are flow diagrams of computer-based procedures.

Figs. 8 and 11 are illustrations of output produced by software.

Figs. 9A-9B are illustrations of computer data.

Detailed Description

Fig. 1 illustrates an Internet analysis system 110 in which a domain names analysis application 112 executes a procedure 1000 (Fig. 2) to collect domain names 114 from a domain name source 116 (step 1010), test the domain names to determine which of the domain names are domain names that correspond to functioning Web sites ("live domain names" 115) (step 1020), and deliver live domain names to a search engine 116 for use in searching the Web (step 1030).

The domain names analysis application collects domain names as follows. The domain name source may include a domain name registry or a root zone server or both. To collect domain names from the domain name registry, the domain names analysis application executes a procedure 2000 (Fig. 3) to submit a request to the domain name registry for a zone file (step 2010), download the requested zone file (step 2020), and extract domain names from the requested zone file (step 2030). In a specific embodiment, the zone file is downloaded by use of a binary transfer procedure known as an FTP transfer. Fig. 11 illustrates an example of a portion of a zone file and extracted domain names. The example is divided into sections. In a typical section, as shown in the example, the first line includes the domain name (in the first column) and its corresponding domain name server (in the last column), and the next line lists the domain name server (in the first column) and its actual IP address (in the last column). If the domain

name has more than one domain name server, the section may include additional lines, each including name and IP address information for another domain name server. After the zone file is downloaded, the domain names are extracted and duplicate domain names are removed.

To collect domain names from the root zone server, the domain names analysis application executes a procedure 3000 (Fig. 4) to request domain name information record by record from the root zone server (step 3010) and extract the domain names from the domain name information (step 3020). In a specific embodiment, domain names are collected from a root zone server as follows. First, the root zone server (e.g., F.Root-Servers.net) is selected and data from the root zone server is directed into a file; the following is a sequence in which the F root server is selected and the F root server is directed to unload all data that ends in "com" into a file called "com.txt".

```
> nslookup
> server f.root-servers.net
> ls com > com.txt
```

Next, a Perl program is executed to extract the domain names from the file in accordance with the principles described above in connection with the zone file example.

Finally, the process, with appropriate changes, is repeated as necessary to collect other domain names. Since different root zone servers are responsible for different domain name extensions (e.g., "com", "net", "edu", "ca", "uk"), collecting a comprehensive list of domain names requires gathering domain name information from multiple root zone servers. In a specific embodiment, other root zone servers are identified by use of a "whois" command. For example, to identify a root zone server that is responsible for "ca" which is the domain name extension for Canada, the following command line is used.

> whois ca-dom

The response generated in this case is "relay.cdnnet.ca". Domain names are gathered from the "relay.cdnnet.ca" server by using a variation of the process described above, which variation uses "relay.cdnnet.ca" in place of "f.root-server.net".

To test a domain name, the domain names analysis application executes a procedure 4000 (Fig. 5) to attempt to acquire the IP address associated with the domain name (step 4010) and, if the IP address is acquired, to attempt, by a request known as an HTTP protocol query, to retrieve an HTML header from a server having the IP address (step 4020). In a specific implementation, a prefix "www." is added to the domain name to form a URL, and the URL is handled much as a typical URL is handled by Web browser software. For example, a

protocol known as telnet is used to attempt to connect to a Web server and retrieve an HTML header. The following command lines illustrate an example for the case of the "uspto.gov" domain name:

> telnet www.uspto.gov 80

> dump

If both attempts 4010, 4020 succeed, the domain name is determined to be a live domain name (step 4030). If either attempt fails, the domain name is determined not to be a live domain name (step 4040). In the case of the attempt to retrieve the HTML header, failure takes the form of a timing out, because the domain names analysis application fails to receive any response. If the Web site returns an error page, it does not qualify as a failure, because the error page includes the HTML header. In such a case, the Web site is functioning, but its contents may be corrupted or may be blocked by security arrangements.

The live domain names are delivered to the search engine as a list that is added to the search engine's list of domain names to be searched for content to be recorded in the search engine's index. In a variation, all or some of the domain names are delivered after being sorted in accordance with a prioritization scheme that takes into account information, gathered from mapping and entity information databases (described below), pertaining to respective entities that are registered as having control over the domain names. For example, the only

domain names delivered may be domain names registered as being under the control of telecommunications companies or companies in a particular city.

Fig. 6 illustrates a variation 200 of the Internet analysis system that allows a user to produce research reports regarding Web activity in sectors of human activity. System 200 includes a mapping database 12 (Fig. 7) that maps URLs or domain names 14 to entities 16 such as people, businesses, or government agencies, as described in more detail below. For example, the mapping database may indicate that any URL that begins with "http://www.uspto.gov" is for a Web page controlled by the U.S. Patent and Trademark Office, or that domain names "elmstdogs.com" and "elmstcats.com" are under the control of a company named Elm Street Pets, Inc. System 200 also includes a Web activity analysis application 202 (described below) and an entity information database 28 (described below) that includes information such as geographic information about entities to which URLs or domain names are mapped in the mapping database.

The mapping database may use a unique identification number ("unique ID"), such as a 9-digit American Business Information ("ABI") number or a DUNNS number, to identify an entity so that other information about the entity can be retrieved from the entity information database or elsewhere by searching under the unique ID. (ABI numbers are sponsored by infoUSA.) For example, unique IDs from the mapping database may be used to search the entity

information database to produce a subset of the mapping database that has records only for entities having a particular characteristic, such as a particular geographic location or between 1000 and 5000 employees.

With respect to the mapping database, where an entity constitutes a portion of another entity, each of the entities may be assigned different unique IDs, and the different unique IDs may be linked in the mapping database to note the relationship among the entities. For example, a company that has offices in different locations may be assigned a unique ID for the company itself and a respective different unique ID for each location. In another example, when two previously unrelated companies merge or one is acquired by the other, each may retain its unique ID and a new, different unique ID may be assigned to the combination of the two companies, or both companies may be assigned the same unique ID.

Information in the mapping database may be derived from information submitted by or on behalf of the entity when a domain name is registered. For example, when the company Elm Street Pets, Inc. registers the domain names "elmstdogs.com" and "elmstcats.com" with a domain name registry, the company associates the domain names with at least enough information, such as name, address, and telephone number information, to allow the domain name registry to bill the company for maintenance of the registration.

The entity may submit information to the mapping database in other ways such as in an on-line questionnaire that feeds the mapping database.

Information in the mapping database may be derived from information provided by an intermediary such as an ISP or an Internet portal. For example, an ISP having a domain name "isp321.com" may have a customer Maple Street Plumbing for which the ISP hosts and administers a home page having a home page address "www.isp321.com/~maplestplumb". In such a case, the ISP may have name, address, and telephone number information for the purpose of billing Maple Street Plumbing for such hosting and administration, and may allow such information along with the home page address to be used to link the home page address to Maple Street Plumbing in the mapping database.

In another example, an Internet portal may allow an entity such as Maple Street Plumbing to create an entry or listing named "Maple Street Plumbing" in a "plumbing" section of a on-line directory maintained by the portal, to allow a user to view home page "www.isp321.com/~maplestplumb" by selecting the entry. In such a case, the Internet portal may allow information in the entry, and perhaps any address and telephone number information submitted by the entity during creation of the entry, to be used to link the home page to Maple Street Plumbing in the mapping database.

The mapping database, the entity information database, and the Web analysis application allow a report such as report 204 (Fig. 8) to be generated that shows, in absolute numbers and as a percentage, how many entities in an industrial sector are registered as controlling one or more Web sites. The entities included in the report may also or instead be limited by geographical area or by any other attribute stored for entities in the entity information database (Figs. 9A-9B illustrate a list of such attributes). To generate the report, the Web analysis application executes a procedure 5000 (Fig. 10) to search the entity information database for entities that match an industrial sector code such as an SIC code or a North American Industrial Classification System ("NAICS") code ("sector entities") (step 5010), determine from the mapping database which of the sector entities are registered as controlling one or more Web sites ("Web entities") (step 5020), and account for each of the sector entities and Web entities in the report (step 5030), such as by presenting quantities for sector entities and Web entities and indicating the number of Web entities as a percentage of the sector entities.

Other reports, such as time based reports, can also be generated by the Web analysis application. For example, the percentage of sector entities that are Web entities can be tracked over time to demonstrate the growth in the number or percentage of entities that are registered as controlling one or more Web sites ("online penetration"). By limiting the entities in the report by entity size (e.g.,

number of employees) or other attribute (e.g., obtained from the entity information database), the report can demonstrate other aspects of Web activity, such as the difference in online penetration among large, medium, and small companies, or which industrial sectors have the most or least online penetration.

The mapping database and applications based on the mapping database may take advantage of a hierarchical organization of Web pages, by treating similarly a mapped page and all pages below the mapped page, such as pages sharing a particular prefix with the mapped page. For example, all pages sharing the prefix "http://www.isp321.com" may be treated as being under the control of an ISP named Global ISP Co.

The mapping database may map an entity to Web pages maintained at different Web sites. For example, Maple Street Plumbing may have a first set of Web pages at the Global ISP Co. site and a second set of Web pages at another ISP's site.

The entity information database may include a database such as EDGAR that includes information about companies.

Information in the mapping database or the entity information database may allow searches to be limited by relative size of entities, such as size in an industry.

One or more of the databases referenced above may be or include a relational database and may have records to which fields may be added readily.

Any of many different types of computer equipment may be used. For example, one or more Intel-based personal computers may be used that run an SQL database on Linux and that programs written in Perl or the C programming language with interfaces to the SQL database.

The technique (i.e., the procedures described above) may be implemented in hardware or software, or a combination of both. In at least some cases, it is advantageous if the technique is implemented in computer programs executing on one or more programmable computers, such as a personal computer running or able to run an operating system such as Unix, Linux, Microsoft Windows 95, 98, or NT, or MacIntosh OS, that each include a processor, a storage medium readable by the processor (including volatile and non-volatile memory and/or storage elements), at least one input device such as a keyboard, and at least one output device. Program code is applied to data entered using the input device to perform the technique described above and to generate output information. The output information is applied to one or more output devices such as a display screen of the computer.

In at least some cases, it is advantageous if each program is implemented in a high level procedural or object-oriented programming language such as Perl, C,

C++, or Java to communicate with a computer system. However, the programs can be implemented in assembly or machine language, if desired. In any case, the language may be a compiled or interpreted language.

In at least some cases, it is advantageous if each such computer program is stored on a storage medium or device, such as ROM or optical or magnetic disc, that is readable by a general or special purpose programmable computer for configuring and operating the computer when the storage medium or device is read by the computer to perform the procedures described in this document. The system may also be considered to be implemented as a computer-readable storage medium, configured with a computer program, where the storage medium so configured causes a computer to operate in a specific and predefined manner.

Other embodiments are within the scope of the following claims. For example, the user may be a human being or a non-human entity such as a computer program or an automated device that may interact with one or more of the databases or one or more of the applications via an application programming interface ("API") or a network message. An on-line information store or multiple databases may serve as the entity information database, which may take the form of any mechanism that provides automated access to information, such as a spreadsheet file or a store of email messages. System 110 may also refer to the mapping and entity information databases before reporting the live domain names

to the search engine. For example, by referring to the mapping and entity information databases, system 110 can retrieve entity information relating to the live domain names, and can sort the live domain names according to the entity information, such as by listing first the live domain names that pertain to an industry that is indicated as being particularly relevant to the search engine or users of the search engine.

What is claimed is:

Claims

1. A method comprising:

acquiring a domain name from a domain name registry;

determining whether the domain name represents any functioning Web site; and

if the domain name is associated with at least one functioning Web site, recording that the domain name represents at least one functioning Web site.

2. The method of claim 1, further comprising:

if it is indicated that the domain name represents at least one functioning Web site, submitting the domain name to a search engine.

3. The method of claim 1, further comprising:

identifying an entity that is registered as having control over the domain name;

determining whether the entity meets a set of prioritization criteria; and

submitting the domain name to a search engine only if the entity meets the set of prioritization criteria.

4. A method comprising:

acquiring a set of criteria;

identifying a set of entities that meet the criteria; and

determining how many entities in the set of entities are registered as having control over at least one Web site.

5. The method of claim 4, comprising:

deriving a statistic from a result of the determination.

6. Computer software, residing on a computer-readable storage medium, comprising a set of instructions for use in a computer system to cause the computer system to:

acquire a domain name from a domain name registry;

determine whether the domain name represents any functioning Web site; and

if the domain name is associated with at least one functioning Web site, record that the domain name represents at least one functioning Web site.

7. Computer software, residing on a computer-readable storage medium, comprising a set of instructions for use in a computer system to cause the computer system to:

acquire a set of criteria;

identify a set of entities that meet the criteria; and

determine how many entities in the set of entities are registered as having control over at least one Web site.

8. A system comprising:

an acquirer that acquires a domain name from a domain name registry;

a determiner that determines whether the domain name represents any functioning Web site; and

a recorder that, if the domain name is associated with at least one functioning Web site, records that the domain name represents at least one functioning Web site.

9. A system comprising:

an acquirer that acquires a set of criteria;

an identifier that identifies a set of entities that meet the criteria; and

a determiner that determines how many entities in the set of entities are registered as having control over at least one Web site.

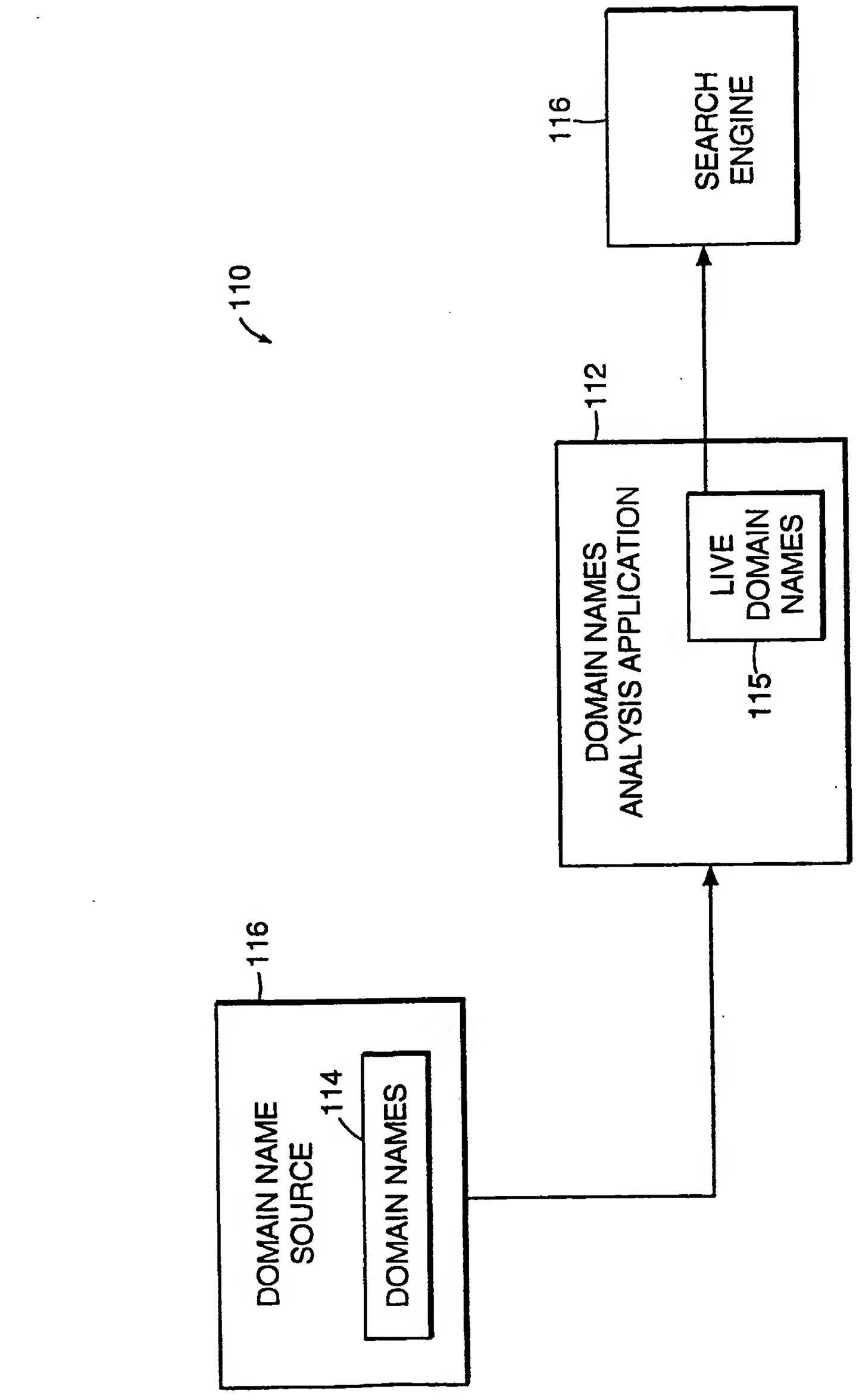

110
116
SEARCH ENGINE
112
DOMAIN NAMES ANALYSIS APPLICATION
LIVE DOMAIN NAMES
115
DOMAIN NAME SOURCE
114
DOMAIN NAMES

FIG. 1

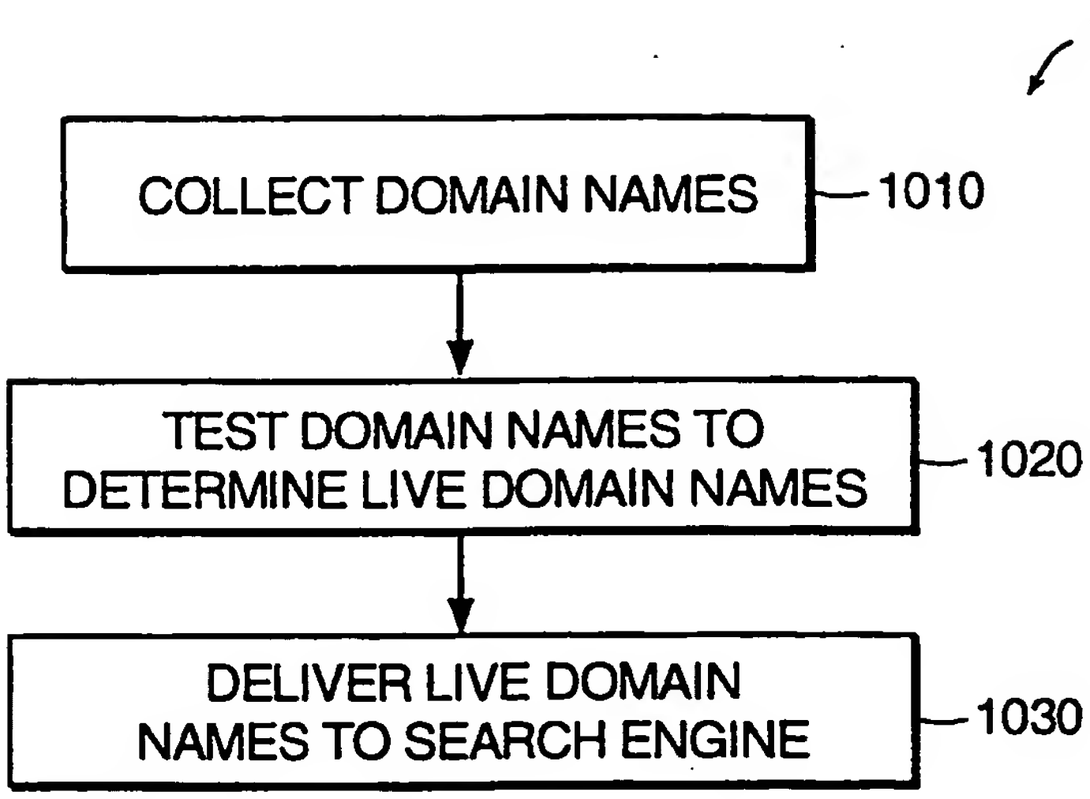


FIG. 2

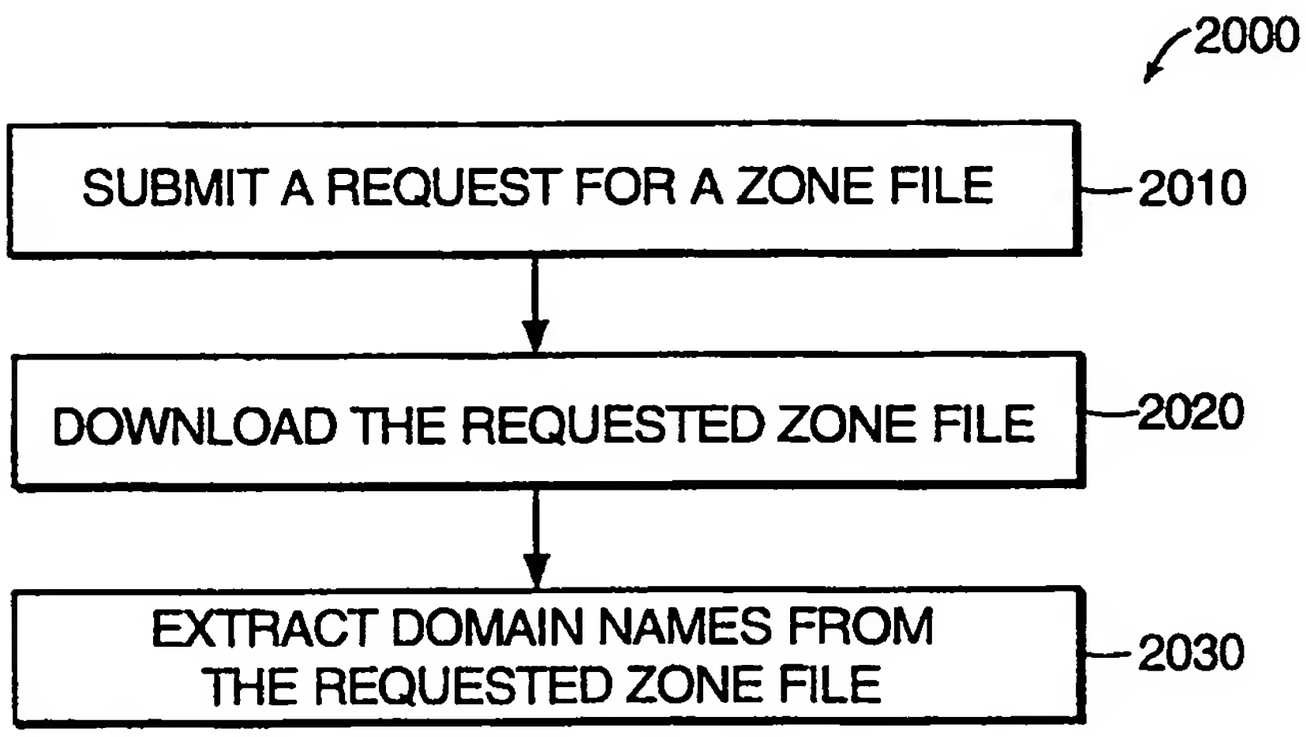


FIG. 3

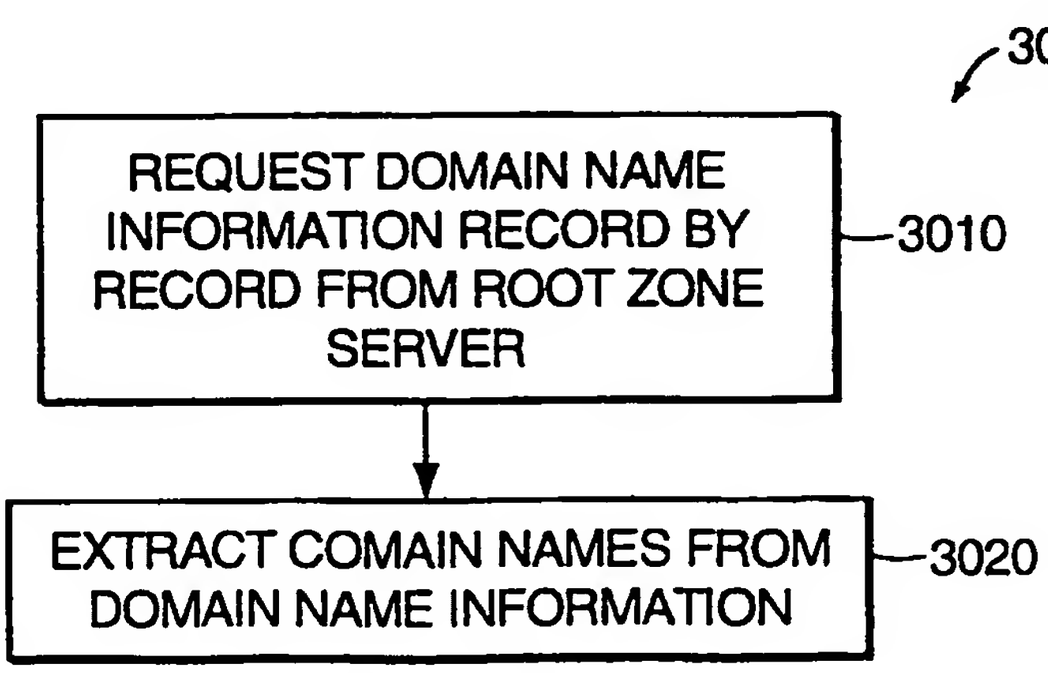


FIG. 4

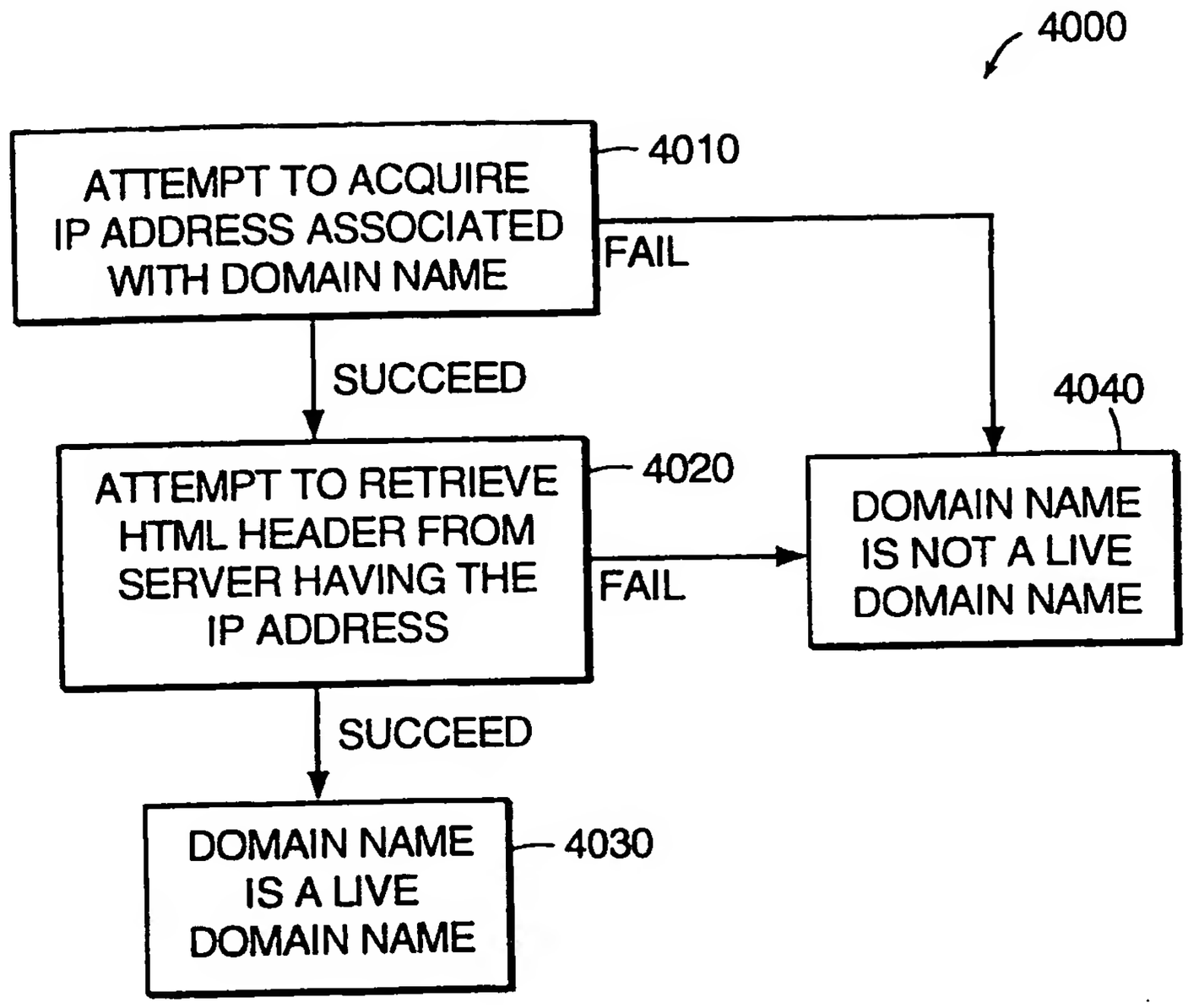

4000
ATTEMPT TO ACQUIRE IP ADDRESS ASSOCIATED WITH DOMAIN NAME
4010
FAIL
SUCCEED
4040
ATTEMPT TO RETRIEVE HTML HEADER FROM SERVER HAVING THE IP ADDRESS
4020
FAIL
DOMAIN NAME IS NOT A LIVE DOMAIN NAME
SUCCEED
DOMAIN NAME IS A LIVE DOMAIN NAME
4030

FIG. 5

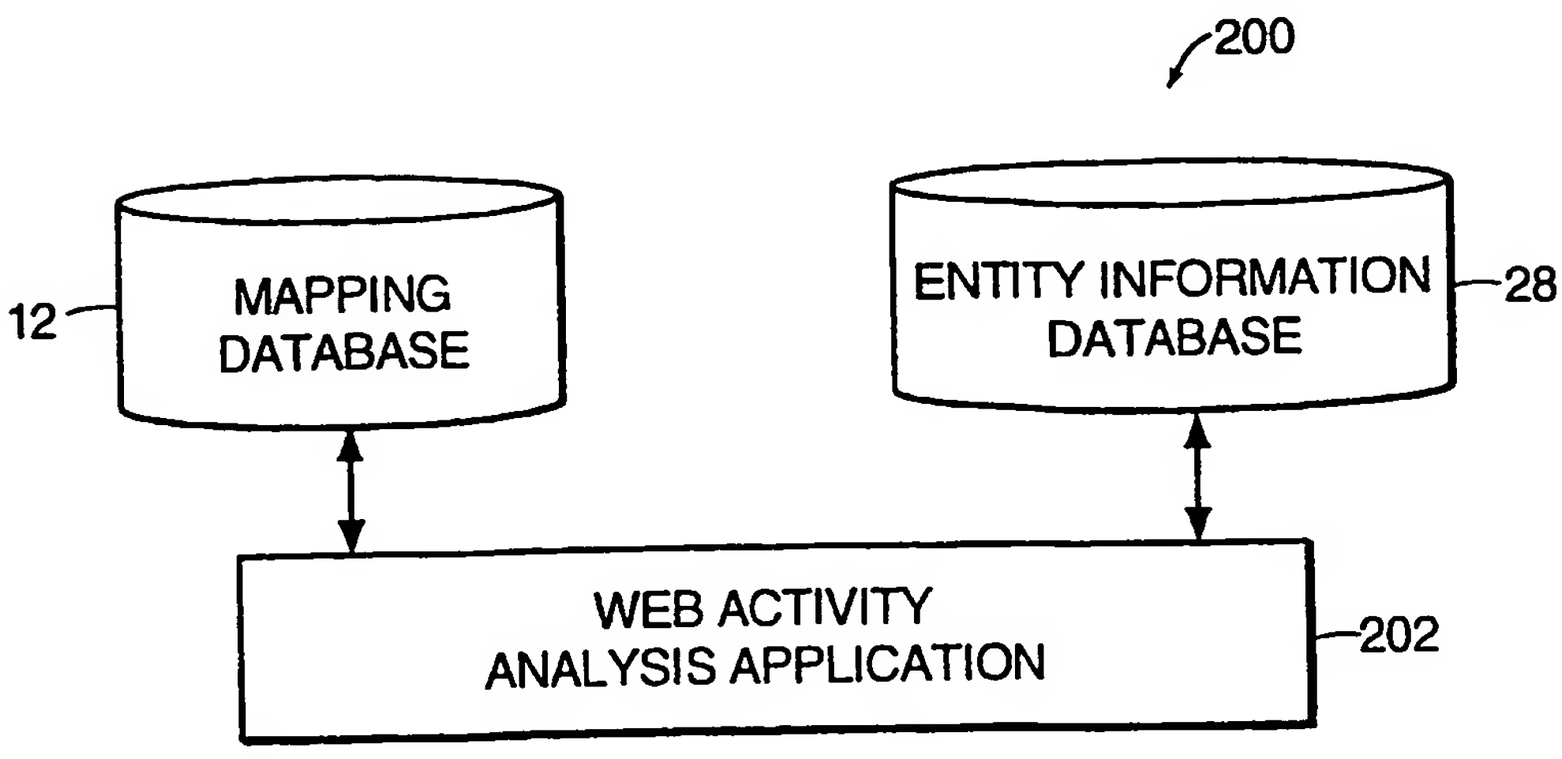

200
12
MAPPING DATABASE
ENTITY INFORMATION DATABASE
28
WEB ACTIVITY ANALYSIS APPLICATION
202

FIG. 6

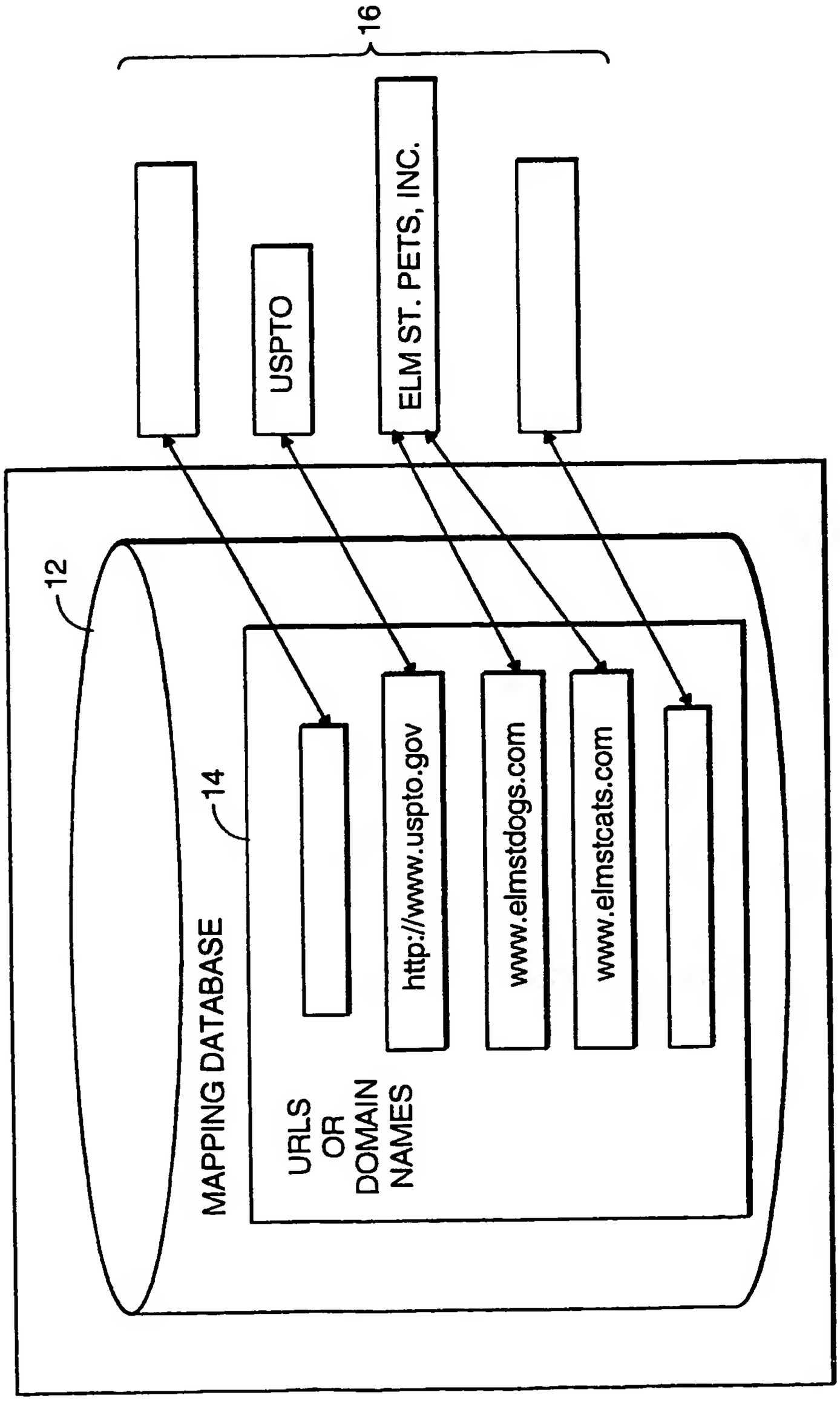

16
USPTO
ELM ST. PETS, INC.
12
14
MAPPING DATABASE
URLS OR DOMAIN NAMES
http://www.uspto.gov
www.elmstdogs.com
www.elmstcats.com

FIG. 7

204

1Atlas

Infol

Home | Research | Register | Help

Search for [ ] in [Industries] [Go]

Use Search to find industries, metropolitan areas, or domain names.

Feedback
Contacts
Company
Credits

Feedback
Become a
Sponser

Company
Profiles

1Atlas

Filter — Use Filter to specify what data to use for a report. Select any combination of industry and geography filters. To reset your filters, click on the circle icon at right.

Industry: All Industries

Geography: Entire US

Report — Use Report to generate a report based on your filters. Select an industry or a geography report at the level of detail represented by the corresponding icon below.

Industry: 2 3 4 5 6 Geography:

NAICS: Low(2) Granularity High(6) USA: Region State MSA:

Line#	Sector (NAICS 2)	Code	Total	Online	Penetration
1	Accommodation and Food Services	72	641,958	140,452	21.88%
2	Administrative and Support and Waste Management and Remediation Services	56	396,087	48,841	12.33%
3	Agriculture, Forestry, Fishing and Hunting	11	61,072	2,222	3.64%
4	Arts, Entertainment, and Recreation	71	166,252	12,117	7.29%
5	Construction	23	739,691	32,302	4.37%
6	Educational Services	61	218,570	15,318	7.01%
7	Finance and Insurance	52	523,761	166,216	31.74%
8	Health Care and Social Assistance	62	1,116,256	38,226	3.42%
9	Information	51	175,612	44,031	25.07%
10	Management of Companies and Enterprises	55	5,194	492	9.47%
11	Manufacturing	31,32,33	479,483	97,478	20.33%
12	Mining	21	28,478	3,777	13.26%
13	Other	99	91,946	4,174	4.54%
14	Other Services (except Public Administration	81	1,391,411	80,802	5.81%
15	Professional, Scientific, and Technical Services	54	1,179,394	121,328	10.29%
16	Public Administration	92	368,484	4,346	1.18%
17	Real Estate and Rental and Leasing	53	469,714	67,704	14.41%
18	Retail Trade	44,45	1,748,844	338,326	19.35%
19	Transportation and Warehousing	48,49	213,938	52,305	24.45%
20	Unclassified	-	232,315	32,009	13.78%
21	Utilities	22	16,626	3,843	23.11%
22	Wholesale Trade	42	625,808	98,431	15.73%
	Totals		10,890,894	1,404,740	12.90%

FIG. 8

- Contact Information
  - Key Executive
  - Contact Last Name
  - Contact First Name
  - Professional Title
  - Title Code
  - Contact Gender Code
  - Admin Contact
  - Admin Contact E-Mail
  - Technical Contact
  - Technical Contact E-Mail
  - Sales Contact Names (1-20)
  - Sales Blind E-Mail Addresses (1-20)
  - Marketing Contact Names (1-20)
  - Marketing Blind E-mail Addresses (1-5)
  - Customer Service Names (1-5)
  - Customer Service E-Mail Addresses (1-20)
  - Investor Relations Name
  - Investor Relations E-Mail Address
  - Purchasing Contact Names (1-5)
  - Purchasing E-Mail Address (1-5)
  - H/R Name
  - H/R E-Mail
  - Executive Names (1-20)
  - Executive Title (1-20)
  - Executive E-mail (1-20)
- Marketing Information
  - Ad Size
  - Population Code
  - Advertise on Web
- Third Party Database Information
  - Individual Firm Code
  - Year of First Appearance
  - Date Added to Database
  - Production Date (MMDDYY)
  - Third Party Number
  - Key Code
- Business Demographics
  - Employee Size Code
  - Sales Volume Code
  - Industry Specific Code
  - Business Status Code
  - Office Size Code
  - Total Employee Size Code
  - Total Output/Sales Code

-]Atlas Database Information
  -]Atlas Number
  - Year of First Appearance
  - Date Added to Database
- Internet Services
  - ISP Number
  - ISP Name
  - Primary Domain Name
  - Secondary Domain Name(S) 1-100
  - Domain Web Servers
  - Domain Web Server Page Count
  - Hosted Indicator
- Internet Demographics
  - Popularity
  - Director
  - MPEG video
  - NetShow
  - QuickTime Movie
  - RealAudio
  - RealMedia
  - Shockwave
- Sales Information
  - E-Commerce Enabled
  - Certified
  - Hours of Operation
  - Brands Carried (1-N)
  - Products Carried (1-N)
  - Current Promotions (1-N)
  - Delivery Available
  - Payment Accepted
  - Language Preference

FIG. 9B

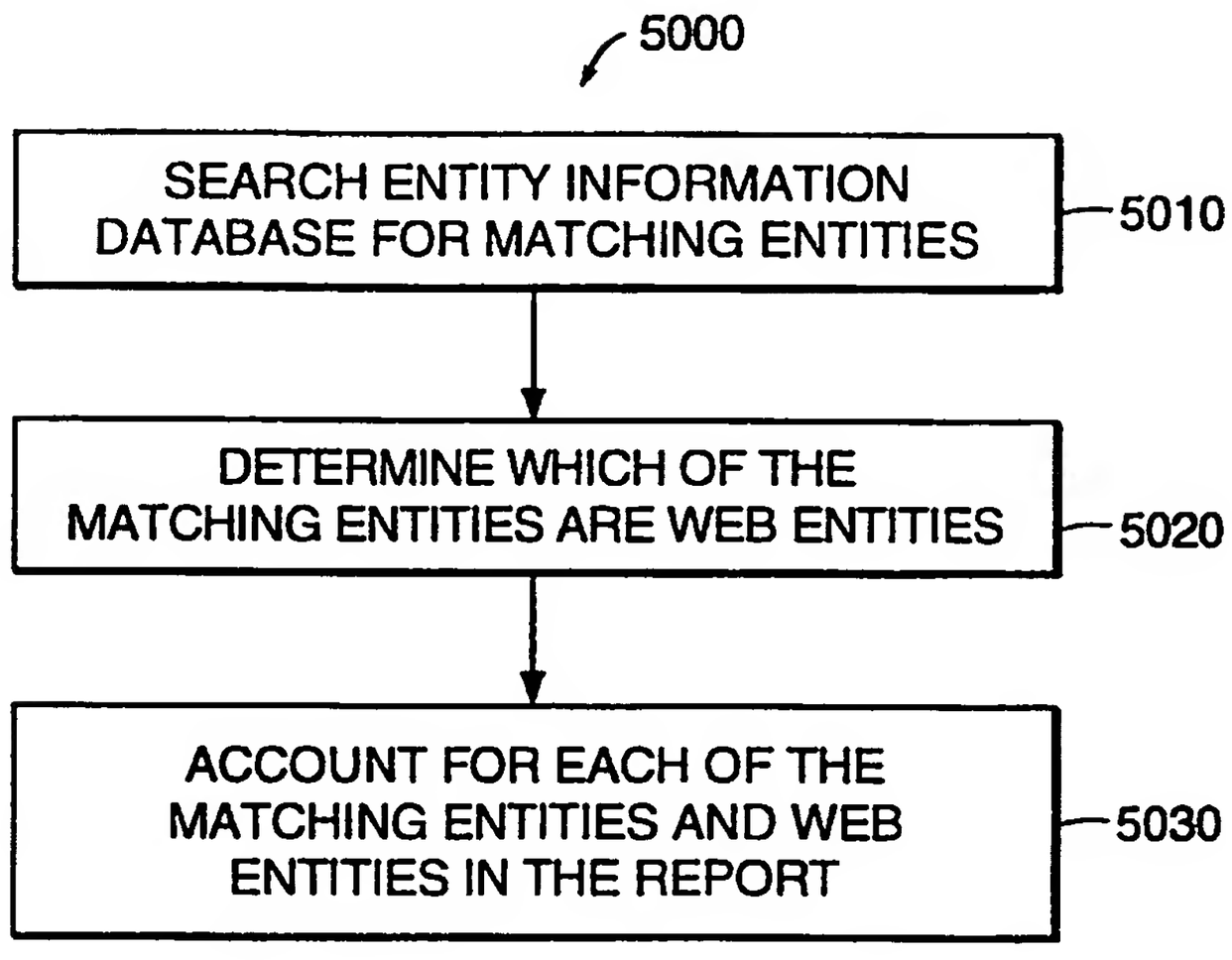

5000
SEARCH ENTITY INFORMATION DATABASE FOR MATCHING ENTITIES
5010
DETERMINE WHICH OF THE MATCHING ENTITIES ARE WEB ENTITIES
5020
ACCOUNT FOR EACH OF THE MATCHING ENTITIES AND WEB ENTITIES IN THE REPORT
5030

FIG. 10

- Company Information(s)
  - Company Name
  - Subsidiary Parent Number
  - Ultimate Parent Number
  - Incorporated as (Company Type)
  - Incorporated in (State)
  - Incorporation Date
  - Years in Business
  - Public Co. Indicator(1)
  - Stock Exchange Code
  - Stock Ticker Symbol
  - Fortune Renlding
  - Stock Price Link
  - News Articles Link
  - Press Release Link
  - Annual Report Link
  - Patents
  - Company Logo
  - Trademarks
- Physical Location
  - Address
  - City(16)
  - State Abbreviation
  - Zip Code
  - Zip+4
  - Phone
  - Fax Phone
  - Carrier Route
  - State Numeric Code
  - Country Code
  - MSA Code
  - Country
  - Latitudes/Longitudes
  - Secondary Address
  - Secondary City
  - Secondary State
  - Secondary Zip Code
  - Number of Locations
- Industry Information
  - Selected SIC Code
  - Franchise Codes
  - Primary Sic
  - Secondary Sic Code #1
  - Secondary Sic Code #2
  - Secondary Sic Code #3
  - Secondary Sic Code #4
  - Additional Industry Codes

FIG. 9A

Zone File Listing:

Domain & Name Server Name	NA	NA	Name Server address
NSP.NET.	172800	NS	NOMAD.NETTEL.COM.
NOMAD.NETTEL.COM.	172800	A	198.245.19.67
REALESTATE.NET.	172800	NS	NS. NEWACCOUNT.COM.
	172800	NS	NS2.NEWACCOUNT.COM.
NCM.NET.	172800	NS	NS.SCRUZ.NET.
NS.SCRUZ.NET.	172800	A	165.227.1,1
NCM.NET.	172800	NS	NS.SCRUZ.NET.
NS2.SCRUZ.NET.	172800	A	165.227.2.10
CHARLESTON.NET.	172800	NS	DNS4.INFOAVE.NET.
	172800	NS	DNS2.INFOAVE.NET.
CARTE.NET.	172800	NS	NS1.GWI.NET.
NS1.GWI.NET.	172800	A	204.120.68.9
CARTE.NET.	172800	NS	DNS.BIHS.NET.
DNS.BIHS.NET.	172800	A	204.96.15.2
APARTMENT.NET.	172800	NS	NS.SYNAPSE-GROUP.COM.
NS.SYNAPSE-GROUP.COM.	172800	A	206.65.140.2
APARTMENT.NET.	172800	NS	NS2.SYNAPSE-GROUP.COM.
NS2.SYNAPSE-GROUP.COM.	172800	A	206.65.140.2
PURE.NET.	172800	NS	NS.CSTONE.NET.
NS.CSTONE.NET.	172800	A	205.197.102.10
ROCKBRIDGE.NET.	172800	NS	NS.ROCKBRIDGE.NET.
NS.ROCKBRIDGE.NET.	172800	A	206.151.4.1
ROCKBRIDGE.NET.	172800	NS	NS.CS.NET.
AUCTION.NET.	172800	NS	NS.NEWACCOUNT.COM.
	172800	NS	NS2.NEWACCOUNT.COM.
NTS.NET.	172800	NS	HUGH.NETTEL.COM.
	172800	NS	NOMAD.NETTEL.COM.
APOLLO.NET.	172800	NS	NS.SOLID.NET.
	172800	NS	NS2.SOLID.NET.
NS2.SOLID.NET.	172800	A	204.119.16.11
PROVENTUM.NET.	172800	NS	DNS.PROVENTUM.DK.
DNS.PROVENTUM.NET.	172800	A	194.19.64.5
PROVENTUM.NET.	172800	NS	DANPOST.UNI-C.DK.
DANPOST.UNI-C.DK.	172800	A	129.142.6.64
A1.NET.	172800	NS	NS1.VIA.AT.
NS1.VIA.AT.	172800	A	194.41.60.10

UNIQUE LISTING OF DOMAINS:

NSP.NET
REALESTATE.NET
NCM.NET
NS.SCRUZ.NET
NCM.NET
CHARLESTON.NET
CARTE.NET
APARTMENT.NET
PURE.NET
ROCKBRIDGE.NET
AUCTION.NET
NTS.NET
APOLLO.NET
PROVENTUM.NET
A1.NET

FIG. 11

INTERNATIONAL SEARCH REPORT

Int. tional Application No
PCT/US 99/18645

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 G06F17/30

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the International search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 97 15018 A (BELL COMMUNICATIONS RES) 24 April 1997 (1997-04-24) abstract page 4, line 5 -page 5, line 3	1,4,6-9
A	WO 97 29414 A (AT & T CORP) 14 August 1997 (1997-08-14) page 2, line 3 -page 3, line 30	1,4,6-9

☐ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the International filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the International filing date but later than the priority date claimed

"T" later document published after the International filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the International search	Date of mailing of the International search report
21 December 1999	11/01/2000

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Katerbau, R

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

Int. tional Application No
PCT/US 99/18645

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9715018 A	24-04-1997	NONE	
WO 9729414 A	14-08-1997	EP 0879448 A	25-11-1998